

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE APRIL 1995		3. REPORT TYPE AND DATES COVERED
4. TITLE AND SUBTITLE MEAN COMPARISONS OF AGE AND GENDER DIFFERENCES ON THE TOTAL SCORE, PUSH-UP SCORE, SIT-UP SCORE, AND 2-MILE RUN SCORE OF THE ARMY PHYSICAL FITNESS TEST			5. FUNDING NUMBERS	
6. AUTHOR(S) MAJOR PHILIP KAHUE MS USA				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) EISENHOWER ARMY MEDICAL CENTER			8. PERFORMING ORGANIZATION REPORT NUMBER 4A-95	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) US ARMY MEDICAL DEPARTMENT CENTER AND SCHOOL BLDG 2841 MCCS HRA US ARMY BAYLOR PGM IN HCA 3151 SCOTT ROAD FORT SAM HOUSTON TEXAS 78234-6135			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The purpose of this retrospective study was to randomly sample APFT scores from soldiers (N=154) assigned to Dwight David Eisenhower Army Medical to see if there were any significant differences in APFT scores based upon age categories and gender.				
14. SUBJECT TERMS MEAN COMPARISONS OF AGE AND GENDER DIFFERENCES			15. NUMBER OF PAGES 41	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

19960911 042

DMC QUALITY INSPECTED

U.S. ARMY - BAYLOR UNIVERSITY
GRADUATE PROGRAM IN HEALTH CARE ADMINISTRATION

GRADUATE MANAGEMENT PROJECT

MEAN COMPARISONS OF AGE AND GENDER DIFFERENCES ON THE
TOTAL SCORE, PUSH-UP SCORE, SIT-UP SCORE, AND 2-MILE
RUN SCORE OF THE ARMY PHYSICAL FITNESS TEST (APFT) OF
ARMY MEDICAL DEPARTMENT SOLDIERS AT DWIGHT DAVID
EISENHOWER ARMY MEDICAL CENTER

SUBMITTED TO
LIEUTENANT COLONEL LEE W. BRIGGS, MS, USA
IN PARTIAL FULFILLMENT
OF REQUIREMENTS FOR THE DEGREE
OF MASTER OF HEALTH CARE ADMINISTRATION

BY
MAJOR PHILIP KAHUE, MS, USA

FORT GORDON, GEORGIA

APRIL 1995

ACKNOWLEDGMENTS

I would like to thank COL John E. Vigna, my preceptor, who provided guidance, understanding and the time for me to complete this project. I would also like to thank CPT Robert Wooldridge and CPT Karen Kirkpatrick, company commanders for Medical Companies A and B, for allowing me to gather data from their training databases to conduct this project. I would also like to thank Dr. Kenn Finstuen, CPT Bruce Reed, and CPT Guy Kiyokawa who worked with me on an earlier project of a similar nature. Their consultation allowed me to come to closure on this project. I would also like to thank my fellow Baylor Resident, MAJ Michael Giovino, for having to listen to me all day, every day, this past year. Finally, I would like to thank my wife Karen and my children for having to put up with me while I finished this project.

ABSTRACT

The United States Army ensures individual readiness through the enforcement of physical fitness standards. These standards are combined in the Army Physical Fitness Test (APFT) which is administered to every soldier at least twice each year. Failure of this test will result in a negative performance evaluation or removal from service. Many soldiers in the youngest age category (17-21 years) cite high standards compared to other age groups.

The purpose of this retrospective study was to randomly sample APFT scores from soldiers ($N = 154$) assigned to Dwight David Eisenhower Army Medical Center to see if there were any significant differences in APFT scores based upon age categories and gender. APFT total scores, push-up scores, sit-up scores, and 2-mile run scores were analyzed with a 2×7 Analysis of Variance experimental design. No differences were found for any effect of gender, nor were any differences found for any interaction effects between gender and age categories for any of the separate APFT events nor the APFT as a whole. However, significant main effect differences did emerge among age categories on the 2-mile run event with $F(6,140) = 4.05$, $p < .001$. Further analysis with Scheffe's mean comparisons method showed that soldiers in the third age category (27-31 years) scored significantly lower than soldiers in the fourth (32-36 years) and fifth (37-41 years) age categories. Results indicate that standards on the 2-mile run may be too high for soldiers in age category 3 (27-31 years). Implications of the findings are discussed.

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	i
ABSTRACT.....	ii
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
Chapter	
1. INTRODUCTION.....	1
Background	
Conditions Which Prompted the Study	
Statement of the Problem or Question	
Literature Review	
Purpose	
2. METHODS AND PROCEDURES.....	10
Study Design	
Validity and Reliability	
Ethical Issues	
3. RESULTS.....	12
4. DISCUSSION.....	29
5. CONCLUSIONS AND RECOMMENDATIONS.....	32
Appendix	
A. DEFINITION OF TERMS.....	34
B. TABLES OF APFT FAILURE RATES IN INTITAL AND FINAL STUDY SAMPLES.....	35
C. SUMMARY OF SUPPORT/NONSUPPORT OF HYPOTHESES.....	36
D. SCHEFFE'S FORMULA FOR COMPARING MULTIPLE MEANS AND CALCULATIONS FOR SIGNIFICANCE.....	38
REFERENCE LIST.....	40

LIST OF TABLES

Table	Page
1. Means and Standard Deviations of Gender and Age Categories for the APFT Total Score.....	13
2. Means and Standard Deviations of Gender and Age Categories for the APFT Push-up Score.....	16
3. Means and Standard Deviations of Gender and Age Categories for the APFT Sit-up Score.....	18
4. Means and Standard Deviations of Gender and Age Categories for the APFT 2-Mile Run Score.....	20
5. Analysis of Variance Source Table and F Ratios for Hypothesis Tests Regarding Mean APFT Total Scores.....	23
6. Analysis of Variance Source Table and F Ratios for Hypothesis Tests Regarding Mean APFT Push-up Scores.....	24
7. Analysis of Variance Source Table and F Ratios for Hypothesis Tests Regarding Mean APFT Sit-up Scores.....	25
8. Analysis of Variance Source Table and F Ratios for Hypothesis Tests Regarding Mean APFT 2-Mile Run Scores.....	27
9. APFT Failure Rates in Initial Study Sample.....	35
10. APFT Failure Rates in Final Study Sample.....	35

LIST OF FIGURES

Figure	Page
1. Mean APFT Total Scores by Age Category.....	14
2. Scatter Plot of APFT Total Score by Age Category...	15
3. Mean APFT Push-up Score by Age Category.....	17
4. Scatter Plot of APFT Push-up Score by Age Category.....	17
5. Mean APFT Sit-up Score by Age Category.....	19
6. Scatter Plot of APFT Sit-up Score by Age Category..	20
7. Mean APFT 2-Mile Run Score by Age Category.....	21
8. Scatter Plot of APFT 2-Mile Run Score by Age Category.....	22
9. Mean APFT Total Score for Male and Female Subjects by Age Category.....	23
10. Mean APFT Push-up Score for Male and Female Subjects by Age Category.....	25
11. Mean APFT Sit-up Score for Male and Female Subjects by Age Category.....	26
12. Mean APFT 2-Mile Run Score for Male and Female Subjects by Age Category.....	27

Mean Comparisons of Age and Gender Differences on the Total
Score, Push-up Score, Sit-up Score and 2-Mile Run Score of the
Army Physical Fitness Test (APFT) of Army Medical Department
Soldiers at Dwight David Eisenhower Army Medical Center

1. INTRODUCTION

It is generally believed by commanders that a high level of physical fitness is necessary to maintain the combat readiness of military troops who must respond to conflicts around the world. Previous research has shown that higher levels of physical fitness lower the cardiovascular risk factors for soldiers (Bielenda et al 1993). Studies have also shown that lower levels of physical fitness also increase potential risk factors for musculoskeletal injuries in soldiers (Knapik et al 1993). While most jobs in the modern military, other than some of the combat arms branches, are no different than those of civilians (Song and Moore 1989), regular exercise and higher levels of fitness improve combat readiness and increase the physical qualities desirable in soldiers (Sharkey 1984). Physically fit soldiers typically have greater levels of self-confidence, mental toughness, and motivation (Pollock and Wilmore 1984). These attributes, regardless of the physical demands of their jobs, contribute to a soldier's performance and the ability to deal with the debilitating stresses of combat.

The United States Army ensures individual readiness through

the enforcement of physical fitness standards. These standards form the Army Physical Fitness Test (APFT) which is a required biannual test for all soldiers.

a. Background

The purpose of the APFT is to insure the maintenance of a base level of physical conditioning that is essential for every soldier in the Army. The APFT is a three-event physical performance test used to assess muscular endurance and cardiorespiratory fitness. The test, in sequence, consists of the push-up event, sit-up event, and 2-mile run. The pushup and sit-up are timed events in which the soldier must attempt as many repetitions as possible in two minutes (Headquarters, Department of the Army Field Manual 21-20).

Adjustments applied to each raw APFT score account for age and physiological differences between men and women. Despite the level of conditioning, there are cardiorespiratory and muscular strength differences between men and women. Without these normalized standards, an equitable comparison cannot be made between the age and gender categories.

The APFT total score consists of the adjusted scores of all three test events. A soldier must score at least sixty points per event and a minimum total score of 180 points to pass the APFT. The maximum for each event is one hundred points. The maximum total score for all three events in a record APFT is three hundred points.

Testing personnel supervise each soldier by counting each

repetition. Proper supervision ensures uniformity in scoring and proper training of supervisors and graders. Field Manual 21-20 gives specific guidance on the description of one repetition of the push-up and sit-up.

The APFT is a valid measure of aerobic capacity and muscular strength and endurance in men and women (Knapik 1989). There exists a high correlation between the 2-mile run and maximum oxygen uptake (VO₂ max) which is a laboratory measure of aerobic fitness and capacity. In a military environment, it is the absolute muscular endurance that is important. The push-up and sit-up provide a measure correlated with muscular strength and endurance (Knapik 1989).

Soldiers must pass a record APFT biannually with a minimum of four months separating record tests if only two tests are given during the year. All soldiers attending courses for 56 days or longer at Army Service Schools or initial training courses must pass the APFT to graduate. Failure to graduate from these courses will prevent promotion to the next grade and may result in eventual removal from the service. Soldiers failing the record APFT for the first time or who do not take the APFT in the required period are administratively flagged. This process prohibits the soldier from any favorable actions such as promotions, formal training, and awards. Performance on this test directly affects a soldier's potential for career progression. Failure of a record APFT will result in a negative annotation on an officer's or noncommissioned officer's annual

evaluation report. Two consecutive failures of this test will result in the processing for elimination of a soldier from the Army.

Soldiers excused from physical fitness training during periods of temporary or permanent medical limitations must maintain a medical profile. Army Regulation 350-15, provides the guidance in administering these medical profiles. A profiling officer designates the medical limitations. Personnel with temporary profiles receive twice the length of the profile (not to exceed ninety days) to train for the APFT. A list of physical activities that are suitable for the profiled soldier accompanies permanent profiles. These alternate events receive either a pass or fail mark.

Army Regulations 623-105 and 623-205, govern officer and noncommissioned officer evaluations, and specify the procedures in noting the pass or failure of an APFT on a soldier's performance evaluation (Headquarters, Department of the Army Regulations 623-105 and 623-205). A notation of failure of an APFT on a performance evaluation will reduce that officer's or noncommissioned officer's chances for promotion or selection for formal Army training.

Modification of the APFT scales in 1987 reflected higher standards for all age categories. For example, before 1987, a male soldier in the youngest age category (17-21 years) received 100 points in each event with 68 push-ups, 69 sit-ups, and a 13 minute 2-mile run. After 1987, the standard for 100 points in

each event rose to 82 push-ups, 92 sit-ups and an 11 minute 54 second 2-mile run. It also created three more age categories for a total of eight. The eight age categories are: 1 (17-21 years); 2 (22-26 years); 3 (27-31 years); 4 (32-36 years); 5 (37-41 years); 6 (42-46 years); 7 (47-51 years); and 8 (52 years and older).

b. Conditions Which Prompted the Study

The perception of equity among the age categories in the APFT scoring is constantly in question. In the February 1994 issue of the Health Services Command Mercury, an article concluded that the APFT allowed older groups to score 100 points in each event with greater ease than for the younger groups. Colonel William Cline writes, "grow as old as possible, you'll have an easier time 'maximizing' the APFT . . . pity the poor 20 year olds who must be Olympic semifinalists to maximize the APFT" (Cline 1994). Although this was not a scientific study, the article underscores the perception of soldiers. Given this evidence, it may be reasoned that the younger age categories may not be properly adjusted for age and gender differences.

Army Medical Department (AMEDD) personnel are not usually subjected to daily physical training to help them meet Army standards. However, all AMEDD soldiers must maintain the same level of physical conditioning as the rest of the Army. A lack of emphasis on this area of training may result in the elimination of health care professionals who maintain a rare and important medical specialty. Unlike other soldiers, the soldiers

in the medical professions are not interchangeable.

c. Statement of the Problem or Question

Are the current age and gender adjustments fair among all age categories for the Army Physical Fitness Test?

d. Literature Review

In 1988, the U.S. Army Physical Fitness School (USAPFS) at Fort Benjamin Harrison, Indiana conducted an active-army physical fitness survey. This study concluded that the senior age groups generally performed well, especially females. It noted that the youngest age group (17-21 years) presented the worst performance with 29 percent of the males and 36 percent of the females failing the APFT (O'Connor et al 1990). This youngest age category was also the only age category that had no test subjects maximize the APFT. The study attributed the poor performance to inadequate leadership in fitness training and low levels of self-motivation. The USAPFS study assumed that the adjustment of raw APFT scores for age and gender categories were accurate. This was the only statistical study in a refereed journal ever published evaluating the Army Physical Fitness Test standards.

In a previous unpublished study conducted by Kahue et al, the researchers, utilizing a 2 x 5 Analysis of Variance (ANOVA) experimental design, found no differences for the main effect of gender, nor were any differences found for any interaction effects between gender and age categories. However, significant main effect differences did emerge among age categories with $F(4,160) = 4.36, p < .003$. Further analysis with Scheffe's mean

comparisons method showed that the youngest age category of 17 to 21 year olds scored significantly lower by 27 points than the second age category (22 to 26 years). It was also interesting to note that the standard deviation for the youngest age category was the smallest among all age categories (from 5 to 17 points smaller) indicating that there was less variability in their physical abilities as a group. Results indicated that standards for the younger age category 1 soldiers may be too high.

However, there were limitations to this study. First, both record and diagnostic APFT scores were used. Second, data subjects were graded on the APFT by different sets of graders. Third, the study was conducted utilizing only the total APFT score. In this study, the three separate event scores along with the total APFT score will be used to determine if a specific event or the total score is responsible for any significant difference detected. Also, the same set of graders will be utilized to grade all test subjects on the APFT, and all APFT results will be taken from record APFT scores.

e. Purpose

The purpose of this study was to assess the mean total and separate event APFT scores from a group of AMEDD soldiers assigned to Dwight David Eisenhower Army Medical Center to determine if there were any significant differences between age categories and gender based on the current Department of the Army standards. The dependent variables in this study are the age and gender adjusted APFT total score, push-up score, sit-up score,

and 2-mile run score. The independent variables are age category and gender. The functional relationships of these variables are stated in hypothesis form as follows:

Ha1) There are significant main effect differences for age category and grand mean APFT total score.

Ho1) There are no significant main effect differences for age category and grand mean APFT total score.

Ha2) There are significant main effect differences for gender and grand mean APFT total score.

Ho2) There are no significant main effect differences for gender and grand mean APFT total score.

Ha3) There are interaction effects between age category and gender in the grand mean APFT total score.

Ho3) There are no interaction effects between age category and gender in the grand mean APFT total score.

Ha4) There are main effect differences for age category and grand mean push-up score.

Ho4) There are no main effect differences for age category and grand mean push-up score.

Ha5) There are main effect differences for gender and grand mean push-up score.

Ho5) There are no main effect differences for gender and grand mean push-up score.

Ha6) There are interaction effects between age category and gender in the grand mean push-up score.

Ho6) There are no interaction effects between age category

and gender in the grand mean push-up score.

Ha7) There are main effect differences for age category and grand mean sit-up score.

Ho7) There are no main effect differences for age category and grand mean sit-up score.

Ha8) There are main effect differences for gender and grand mean sit-up score.

Ho8) There are no main effect differences for gender and grand mean sit-up score.

Ha9) There are interaction effects between age category and gender in the grand mean sit-up score.

Ho9) There are no interaction effects between age category and gender in the grand mean sit-up score.

Ha10) There are main effect differences for age category and grand mean 2-mile run score.

Ho10) There are no main effect differences for age category and grand mean 2-mile run score.

Ha11) There are main effect differences for gender and grand mean 2-mile run score.

Ho11) There are no main effect differences for gender and grand mean 2-mile run score.

Ha12) There are interaction effects between age category and gender in the grand mean 2-mile run score.

Ho12) There are no interaction effects between age category and gender in the grand mean 2-mile run score.

If there are significant main effect differences for either

age category or gender, then current age and gender adjustments for the APFT should be investigated by the U.S. Army Training and Doctrine Command who is the approving authority for the APFT standards.

2. METHODS AND PROCEDURES

The term "soldiers" in this study refers to officers, warrant officers, and enlisted personnel. The initial study sample was derived from APFT scores of 796 soldiers assigned to Dwight David Eisenhower Army Medical Center located at Fort Gordon, Georgia. All soldiers in the sample were active duty and did not possess a medical profile. This was necessary to obtain a complete three-event APFT score. The scores were derived from the October 1994 APFT test schedule administered at Dwight David Eisenhower Army Medical Center. Personnel grading the participants remained throughout the testing cycle to increase the reliability of the test procedures. It was assumed that personnel who graded the participants taking the APFT used equal judgment in applying the standards set forth in Field Manual 21-20. With this assumption, it is believed that these test procedures minimized the respondent, situation (environment), and experimenter (measurer) sources of error.

The sample cell size for each age and gender category were maximized based on the age or gender category with the least amount of subjects necessary to maintain rigor of the test. This permitted maximizing the experimental variance of the research through increasing the sample group sizes. In order to ensure

equal numbers for each gender and age category, the oldest age category (52+ years) had to be eliminated from the study due to insufficient numbers of subjects. The final size of the sample was selected using random selection with the exception of one group which contained only 11 feasible numbers. This group represents females in the seventh age category, and is fully represented, hence the final sample size of 11 for each cell. To achieve equal cell sizes, observations were randomly narrowed to 11 for each of the fourteen age and gender categories studied, resulting in a final sample size of 154.

a. Study Design

Data sources from which these scores were collected include the training databases of Medical Company A and Medical Company B of Dwight David Eisenhower Army Medical Center.

Collection of data was conducted in a retrospective manner. Data collected on each individual included the score for each of the three events, and the APFT total score. Using the three event scores and total score for each gender and age category, a 2 x 7 multivariate data design was developed to allow for mean comparisons between two levels of gender and multiple age categories simultaneously.

b. Validity and Reliability

Because of the standardized testing procedures and consistency of graders in grading the APFT, it was assumed that the measures taken were both reliable and valid. In order to control extraneous variables, known variables (such as profile

persons) were controlled through exclusion.

Internal validity for this experiment can only be assured to the extent that the retrospective data collection was accurate. From the standpoint of this study alone, internal validity was assumed, relative to retrospective data analysis.

c. Ethical Issues

Throughout this study, anonymity and confidentiality was maintained. Data collection consisted of APFT scores, age, and gender. Names were omitted from data collection making it impossible for a connection to be made or inferred between the information in this study and any individual. For this reason, obtaining informed consent was not necessary.

3. THE RESULTS

Separate two factor ANOVAs were utilized to assess the effects of gender and age category upon the dependent variables grand mean APFT total score, grand mean push-up score, grand mean sit-up score, and grand mean 2-mile run score. Further comparisons of age categories with grand mean 2-mile run score were obtained with the use of Scheffe's method of multiple comparison of means.

Table 1 presents the descriptive statistics for the gender and age categories of the study sample for the APFT total score.

TABLE 1. Means and Standard Deviations of Gender and Age Categories for the APFT Total Score

Age Category	Yrs	N	Mean		Mean		<u>Overall</u>	
			Male	SD	Female	SD	Mean	SD
1	17-21	22	219.00	50.33	197.55	39.20	208.27	45.37
2	22-26	22	211.36	36.91	238.91	32.30	225.14	36.66
3	27-31	22	212.82	46.10	213.73	46.18	213.27	45.03
4	32-36	22	229.64	20.35	249.82	22.20	239.73	23.21
5	37-41	22	233.45	28.15	239.27	35.98	236.36	31.67
6	42-46	22	233.36	30.35	221.27	27.85	227.32	29.09
7	47-51	22	242.27	32.58	218.36	51.84	230.32	43.99
Overall		154	225.99	36.49	225.56	39.91	225.77	38.11

Overall results indicate that mean scores for gender across all age categories are very close, as are the standard deviations (225.99 points, 36.49 standard deviation points for male, and 225.56 points, 39.91 standard deviation points for female). Mean scores for age category 1 were the lowest, with a noticeable rise for age category 2, then another drop at age category 3, then a very sharp rise to age category 4, and then a steady decline in scores for the older age categories. Figure 1 graphically depicts the mean APFT Total Scores for the age categories.

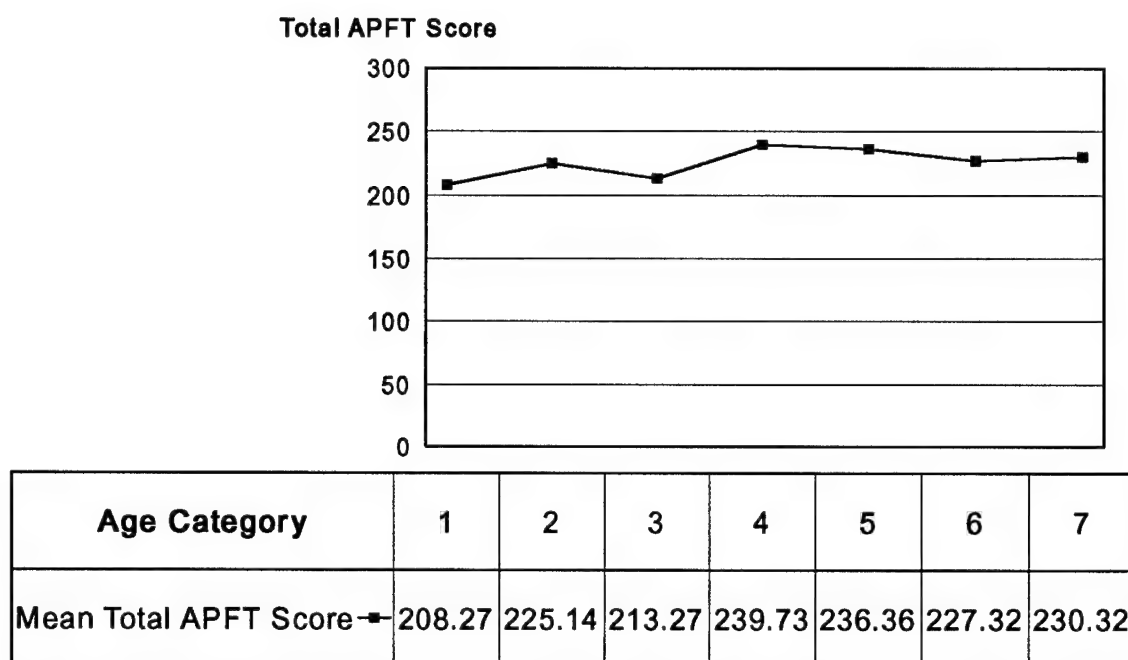


Figure 1. Mean APFT Total Scores by Age Category

It is of particular interest that the standard deviations were larger for age categories 1 (17-21 years), 3 (27-31 years), and 7 (47-51 years). These age categories experienced the greatest numbers of failures with 5, 2, and 3 respectively, thereby increasing the standard deviations.

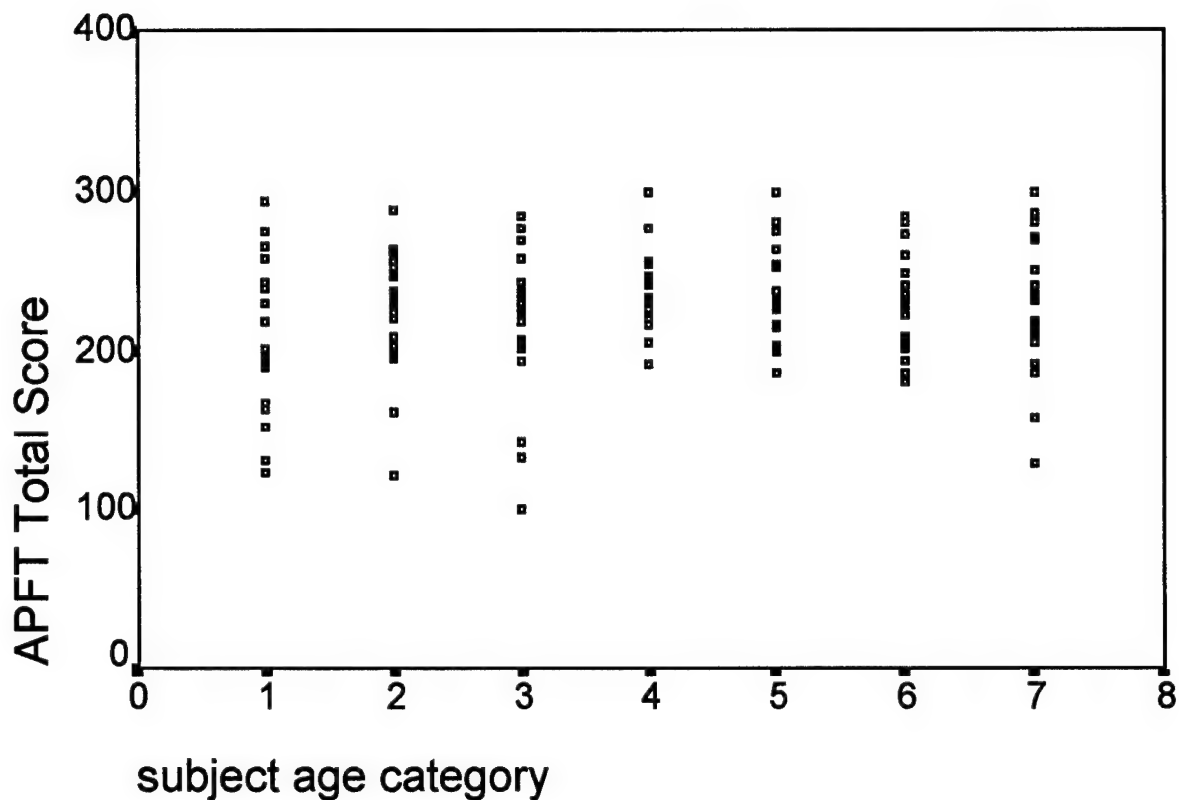


Figure 2. Scatter Plot of APFT Total Score by Age Category

From a scatter plot of the APFT total scores, we can see the dispersion of scores within each age category to account for the variation. Each dot represents one test subject. The rate of failures in these age categories was representative of the failure rates in the initial study population (See Tables 9 and 10 located at Appendix B).

Table 2 presents the descriptive statistics for the gender and age categories of the study sample for the APFT push-up score.

TABLE 2. Means and Standard Deviations of Gender and Age Categories for the APFT Push-up Score

Age Category	Yrs	N	Mean		Mean		Overall	
			Male	SD	Female	SD	Mean	SD
1	17-21	22	76.00	14.69	66.36	8.73	71.18	12.78
2	22-26	22	73.45	12.22	81.09	11.49	77.27	12.22
3	27-31	22	77.91	13.73	72.82	9.36	75.36	11.76
4	32-36	22	75.91	11.75	77.36	9.43	76.64	10.42
5	37-41	22	72.91	10.56	73.27	13.64	73.09	11.90
6	42-46	22	75.36	12.23	69.73	6.75	72.55	10.07
7	47-51	22	81.91	12.01	74.00	12.20	77.96	12.49
Overall		154	76.21	12.34	73.52	10.99	74.86	11.72

Again, overall results indicate that mean scores for gender across all age categories are very close, as are the standard deviations (76.21 points, 12.34 standard deviation points for male, and 73.52 points, 10.99 standard deviation points for female). As before, mean scores for age category 1 were the lowest, with a noticeable rise for age category 2, then a drop to age category 3, then a sharp rise to age category 4, with a steady decline through age categories 5 and 6, and a rise to age category 7. Figure 3 graphically depicts the mean APFT push-up scores for the different age categories.

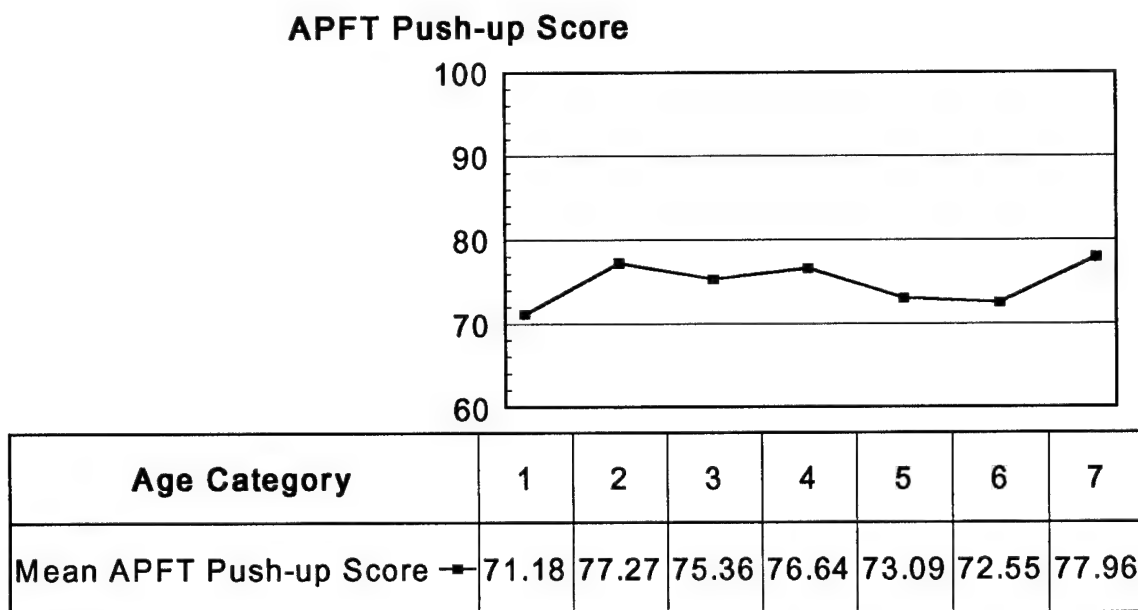


Figure 3. Mean APFT Push-up Score by Age Category

A scatter plot of the APFT push-up scores is unremarkable.

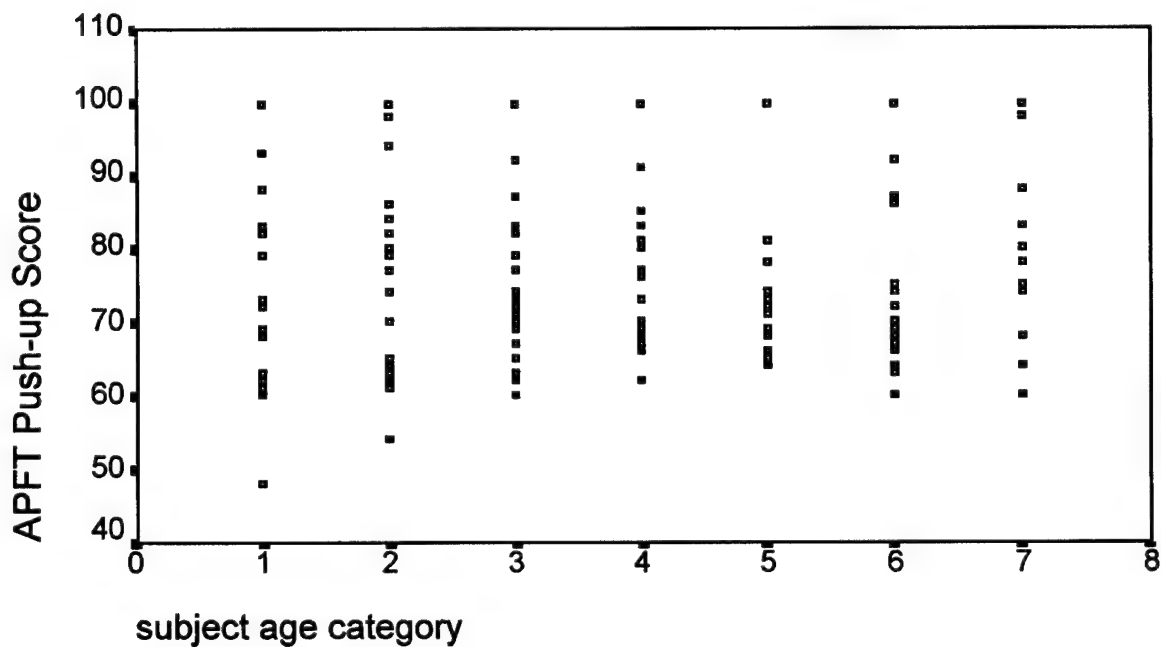


Figure 4. Scatter Plot of APFT Push-up Score by Age Category

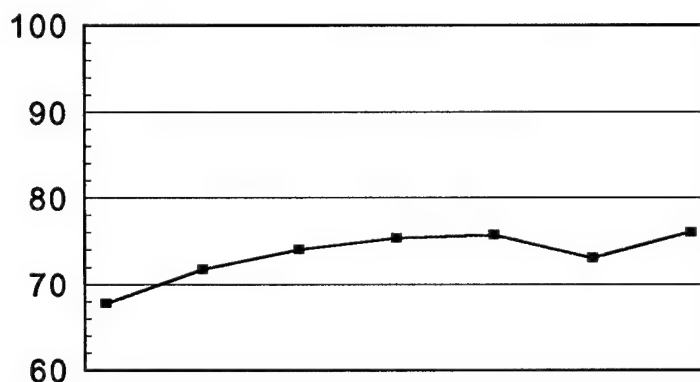
Table 3 presents the descriptive statistics for gender and age categories of the study sample for the APFT sit-up score.

TABLE 3. Means and Standard Deviations of Gender and Age Categories for the APFT Sit-up Score

Age Category	Yrs	N	Mean		Mean		Overall	
			Male	SD	Female	SD	Mean	SD
1	17-21	22	68.45	19.95	67.09	11.97	67.77	16.07
2	22-26	22	68.55	7.10	75.00	8.94	71.77	8.55
3	27-31	22	74.00	11.26	74.09	15.51	74.05	13.23
4	32-36	22	72.73	11.70	78.09	12.61	75.41	12.18
5	37-41	22	75.91	10.56	75.64	13.65	75.77	12.46
6	42-46	22	73.55	10.23	72.64	9.50	73.09	9.65
7	47-51	22	76.64	14.54	75.55	15.16	76.09	14.51
Overall		154	72.83	12.76	74.01	12.61	73.42	12.66

Again, overall results indicate that mean scores for gender across all age categories are very close, as are the standard deviations (72.83 points, 12.76 standard deviation points for male, and 74.01 points, 12.61 standard deviation points for female). From Figure 5 we can see that the general trend of mean scores is slightly different for the sit-up scores.

APFT Sit-up Score



Age Category	1	2	3	4	5	6	7
Mean APFT Sit-up Score	67.77	71.77	74.05	75.41	75.77	73.09	76.09

Figure 5. Mean APFT Sit-up Score by Age Category

Again age category 1 (17-21 years) had the lowest mean score. However, unlike the other APFT scores, the sit-up scores show a gradual increase as age increased, with the exception of age category 6 (42-46 years). A scatter plot (Figure 6) of APFT sit-up scores shows that age categories 1 (17-21 years), 3 (27-31 years), and 7 (47-51 years) experienced the greatest dispersion accounting for their higher standard deviation points.

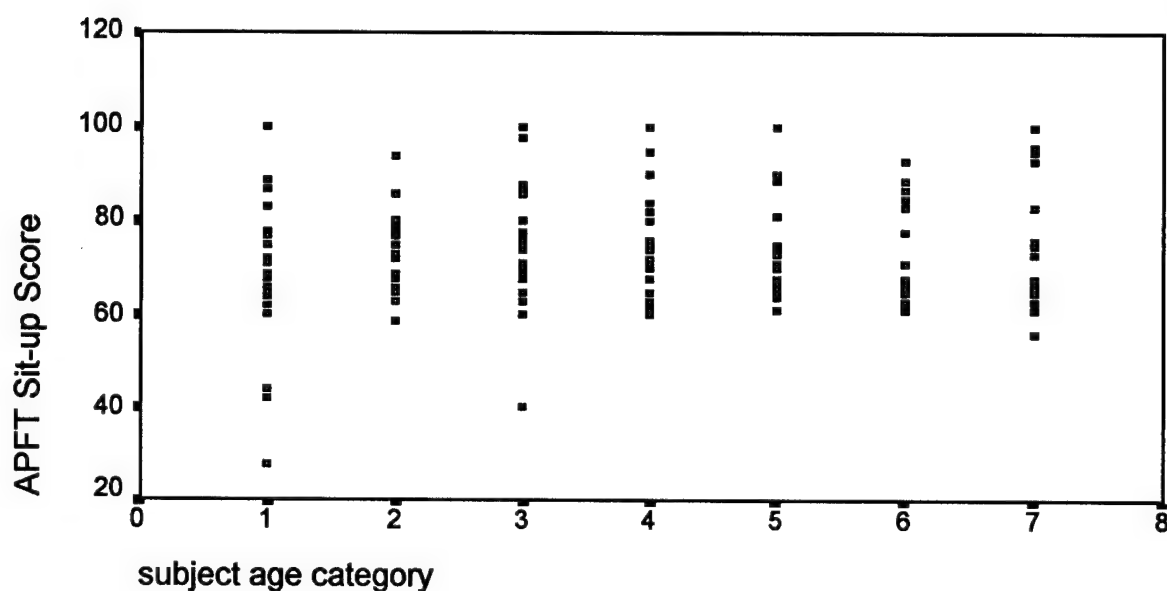


Figure 6. Scatter Plot of APFT Sit-up Score by Age Category

Table 4 presents the descriptive statistics for the gender and age categories of the study sample for the APFT 2-mile run score.

TABLE 4. Means and Standard Deviations of Gender and Age Categories for the APFT 2-Mile Run Score

Age Category	Yrs	N	Mean Male	SD	Mean Female	SD	Overall Mean	SD
1	17-21	22	74.55	24.51	64.09	26.31	69.32	25.38
2	22-26	22	69.36	25.22	82.82	17.90	76.09	22.42
3	27-31	22	60.91	30.95	66.82	25.39	63.86	27.79
4	32-36	22	81.00	7.29	94.36	10.51	87.68	11.17
5	37-41	22	84.64	9.73	90.36	15.79	87.50	13.13
6	42-46	22	84.45	13.72	78.91	21.13	81.68	17.62
7	47-51	22	83.73	13.36	68.82	29.61	76.27	23.68
Overall		154	76.95	20.70	78.03	23.71	77.49	22.19

Overall results indicate that mean scores for gender across all age categories are very close, as are the standard deviations (76.95 points, 20.70 standard deviation points for male, and 78.03 points, 23.71 standard deviation points for female). From Figure 7 we can see that for the first time age category 1 did not have the lowest score among all age categories.

APFT 2-Mile Run Score

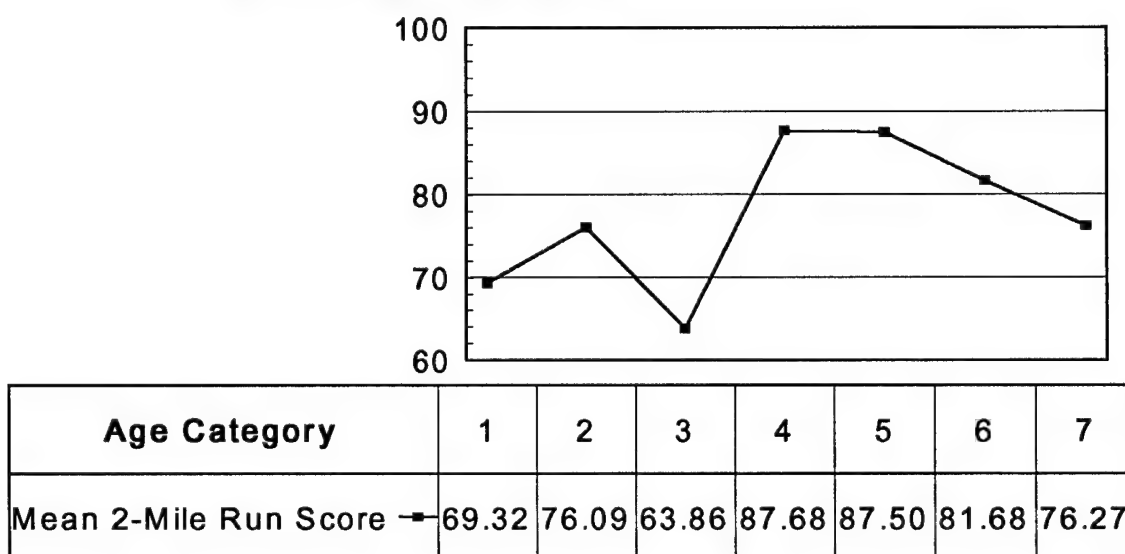


Figure 7. Mean APFT 2-Mile Run Score by Age Category

Age category 3 (27-31 years) actually had the lowest mean score. However, the trend of high and low mean scores appears to follow the total and push-up APFT scores. From a scatter plot (Figure 8) of the 2-mile run scores we can see the dispersion of scores among the age categories.

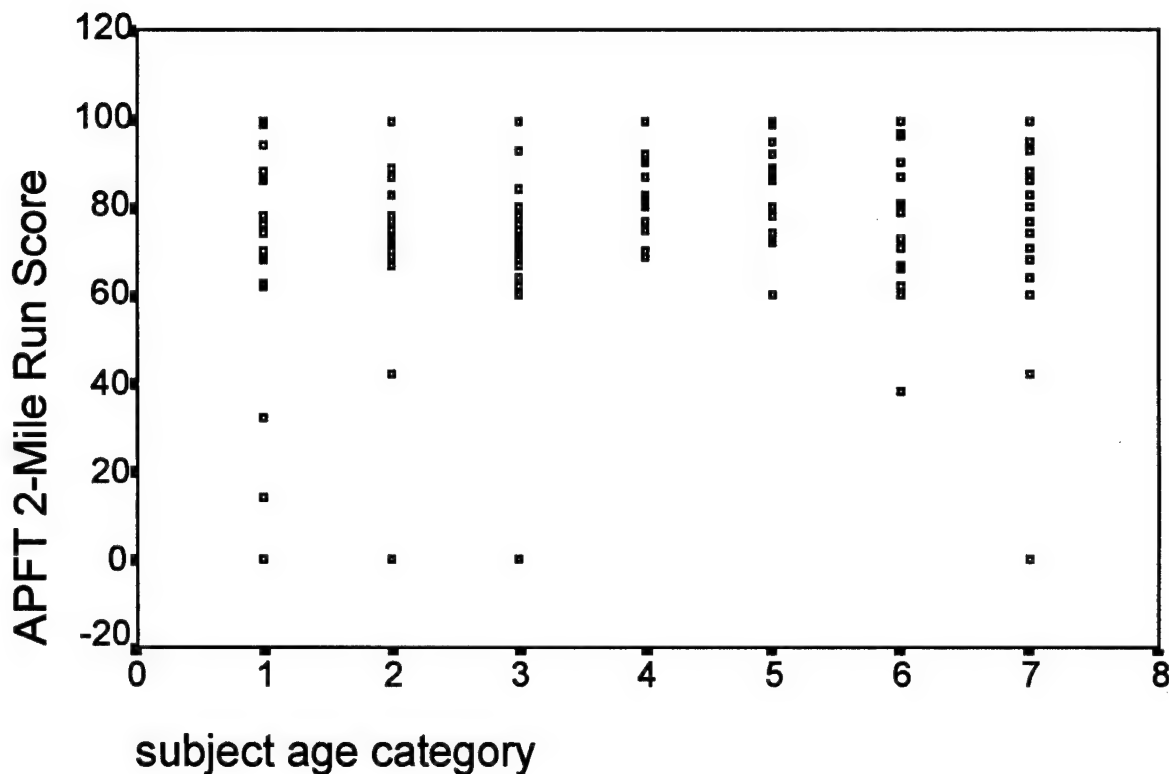


Figure 8. Scatter Plot of APFT 2-Mile Run Score by Age Category

The smaller standard deviation points for age categories 4 and 5 are visible.

A separate 2 x 7 unweighted means ANOVA for equal subject groups for each of the separate APFT events along with the APFT total score was conducted to determine whether statistically significant differences existed among the gender means and among the age category means, and to detect the presence of interaction effects. Results for these hypotheses are displayed in Tables 5 through 8 that follow and in Appendix C.

TABLE 5. Analysis of Variance Source Table and *F* Ratios for Hypothesis Tests Regarding Mean APFT Total Scores

Source of Variation	Sum of Squares	df	Mean Square	<i>F</i>	<i>p</i>
Gender	7.07	1	7.07	.01	.943
Age Categories	17442.73	6	2907.12	2.12	.054
Interaction	13076.70	6	2179.45	1.59	.154
Residual (within)	191718.55	140	1369.42		
Total	222245.05	153	1452.58		

No significant main effect differences were noted for gender and grand mean APFT total score with $F < .01NS$.

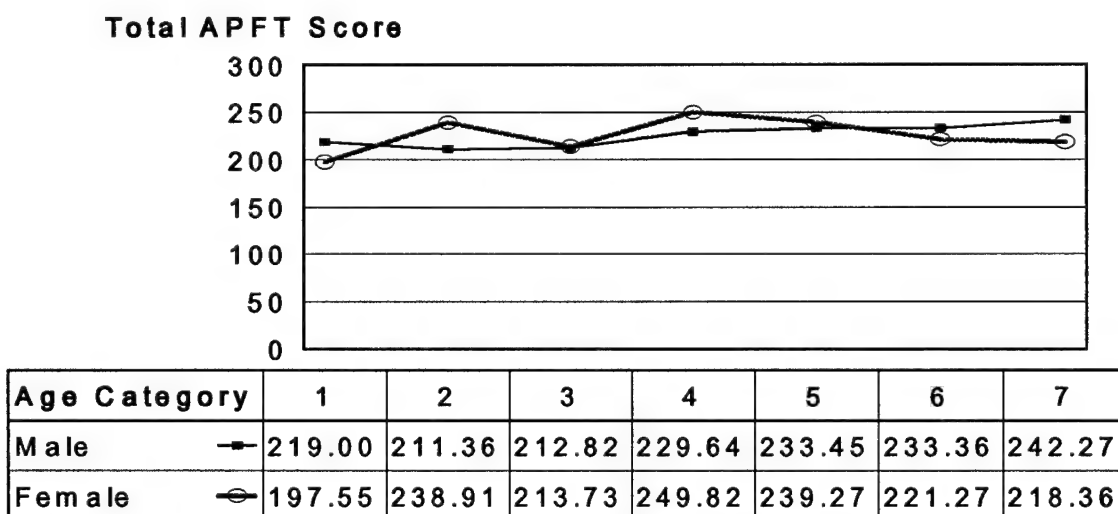


Figure 9. Mean APFT Total Score for Male and Female Subjects by Age Category

A graphical representation of the mean scores for males and females for the APFT total score show that the lines cross indicating that higher scores are not gender specific.

Significance for main effect differences of age category just missed the alpha level set at .05 with $F = 2.12$ NS , $p < .054$, based on 6 and 140 degrees of freedom. This means that the probability that the results obtained could be due to chance exceeds the 95% confidence limit. The absence of any statistically significant interaction indicates that the main effect for age category may be interpreted independently of that of gender and does not require further qualification.

TABLE 6. Analysis of Variance Source Table and F Ratios for Hypothesis Tests Regarding Mean APFT Push-up Scores

Source of Variation	Sum of Squares	df	Mean Square	F	p
Gender	278.24	1	278.24	2.09	.150
Age Categories	898.09	6	149.68	1.13	.351
Interaction	1226.90	6	204.48	1.54	.170
Residual (within)	18622.91	140	133.02		
Total	21026.14	153	137.43		

No significant main effect differences were noted for gender nor age category and grand mean APFT push-up score with $F < 2.09$ NS and $F < 1.13$ NS respectively.

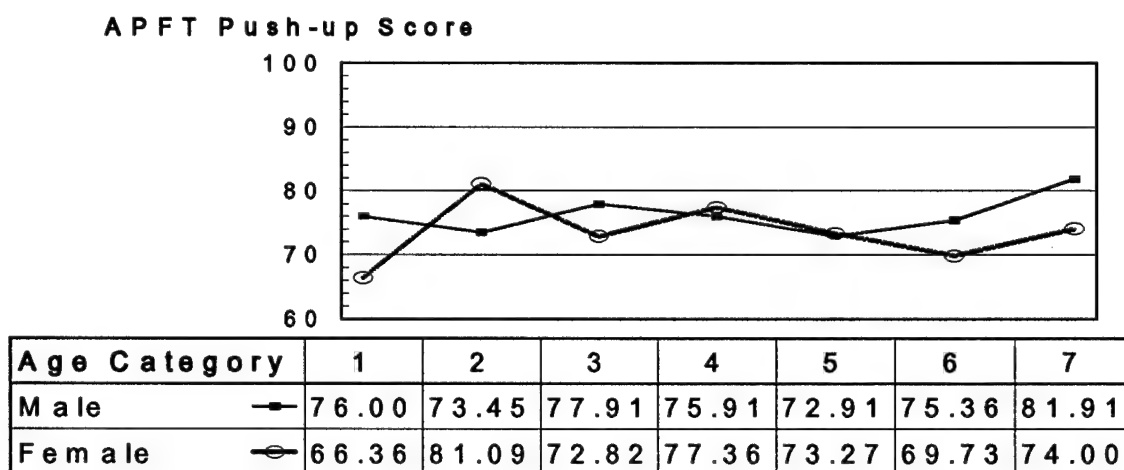


Figure 10. Mean APFT Push-up Score for Male and Female Subjects by Age Category

A graphical representation of the mean scores for males and females for the APFT push-up score show that the lines cross, again indicating that higher scores were not gender specific. The absence of any statistically significant interaction indicates that the main effect for age category may be interpreted independently of that of gender and does not require further qualification.

TABLE 7. Analysis of Variance Source Table and *F* Ratios for Hypothesis Tests Regarding Mean APFT Sit-up Scores

Source of Variation	Sum of Squares	df	Mean Square	<i>F</i>	<i>p</i>
Gender	53.77	1	53.77	.33	.568
Age Categories	1138.07	6	189.68	1.16	.333
Interaction	355.36	6	59.23	.36	.902
Residual (within)	22968.36	140	164.06		
Total	24515.57	153	160.23		

No significant main effect differences were noted for gender nor age category and grand mean APFT sit-up score with $F < .33NS$ and $F < 1.16NS$ respectively.

APFT Sit-up Score

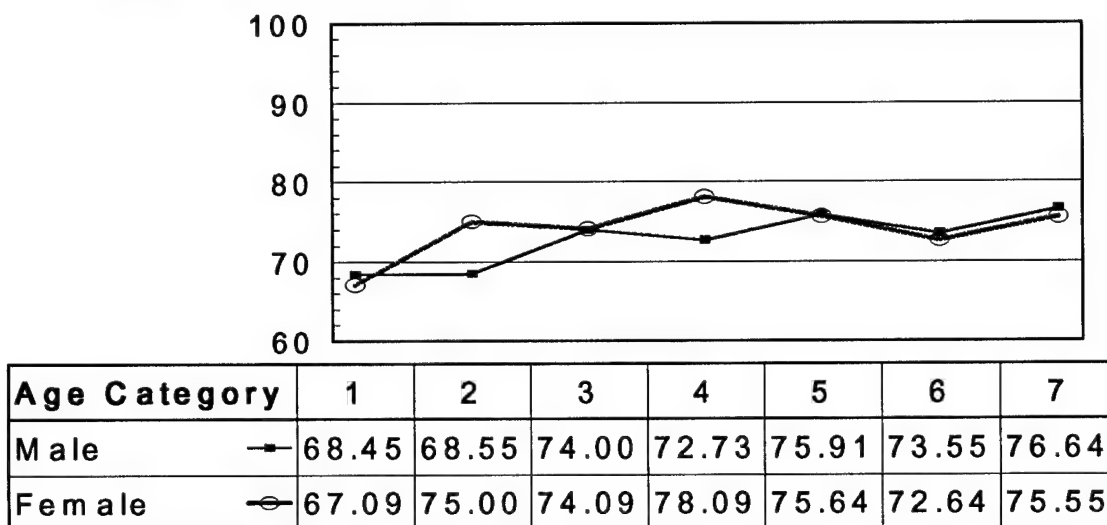


Figure 11. Mean APFT Sit-up Score for Male and Female Subjects by Age Category

Again a graphical representation of the mean scores for males and females for the APFT sit-up score show that the lines cross indicating that higher scores are not gender specific. The absence of any statistically significant interaction again indicates that the main effect for age category may be interpreted independently of that of gender and does not require further qualification.

TABLE 8. Analysis of Variance Source Table and *F* Ratios
for Hypothesis Tests Regarding Mean APFT 2-Mile Run Scores

Source of Variation	Sum of Squares	df	Mean Square	<i>F</i>	<i>p</i>
Gender	44.73	1	44.73	.10	.748
Age Categories	10505.88	6	1750.98	4.05	.001
Interaction	4298.40	6	716.40	1.66	.136
Residual (within)	60495.46	140	432.11		
Total	75344.47	153	492.45		

No significant main effect differences were noted for gender and grand mean APFT 2-mile run score with $F < .10$ NS.

APFT 2-Mile Run Score

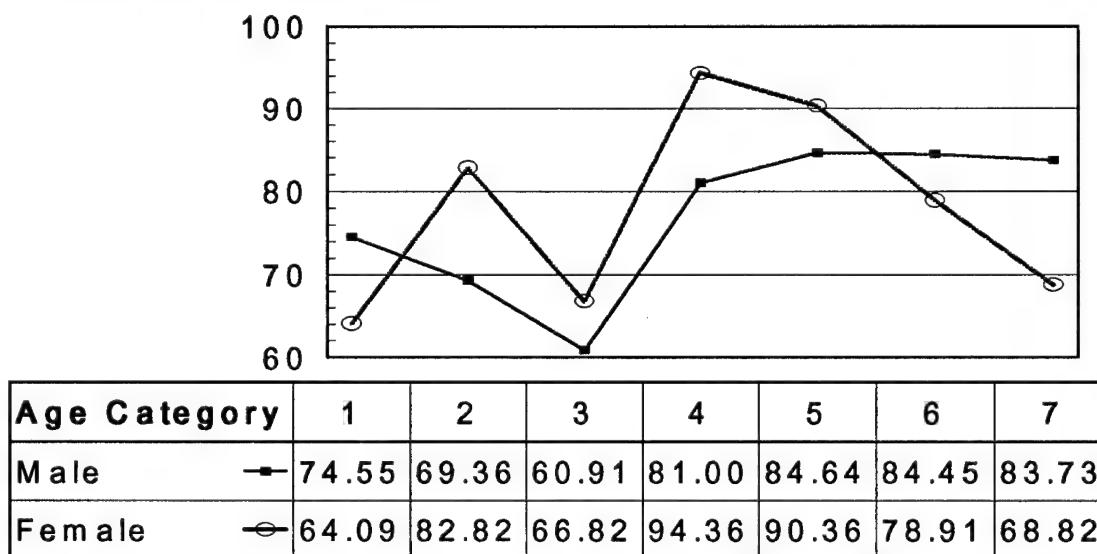


Figure 12. Mean APFT 2-Mile Run Score for Male and Female Subjects by Age Category

A graphical representation of the mean scores for males and females for the APFT 2-mile run score show that the lines cross

indicating that higher scores are not gender specific.

In support of the tenth hypothesis, the main effect for differences in age category means emerged as highly significant, with a F ratio of 4.05, based on 6 and 140 degrees of freedom. This finding was interpreted as support for the expected differences between age categories. The absence of any statistically significant interaction indicates that the main effect for age category may be interpreted independently of that of gender and does not require further qualification.

To further amplify age category mean differences in the APFT 2-mile run score, Scheffe's method for comparing means was used. The alpha level was set at .05, with 6 and 140 degrees of freedom, and the standard error of the difference was based upon the residual mean square between groups. A critical difference value of 22.25 was rendered ($p < .05$). Twenty-one pairwise comparisons were possible among the APFT 2-mile run scores for the seven age category groups. Of those comparisons, two significant difference of means occurred between age categories 3 (27-31 years) and 4 (32-36 years), and between age categories 3 (27-31 years) and 5 (37-41 years). The fourth age group's mean (87.68) exceeded the third age group's mean (63.86) by 23.82 APFT score points, and the fifth age group's mean (87.50) exceeded the third age group's mean (63.86) by 23.64 APFT score points. Scheffe's formula and calculations for significance are located in Appendix D.

4. DISCUSSION

The results of this study suggest that the current gender adjustment for all APFT events across all age categories appears to be equitable. However, the age adjustment for the 2-mile run does not appear to be uniform across all age categories. The results indicate that the current age adjustment for 27-31 year olds (age category 3) may be too high on the 2-mile run.

Scheffe's method was used in this study to test for statistical significance in difference of means. This method is the most conservative for comparing multiple means, however, it protects the alpha level best (Dowdy and Wearden 1991). Other methods build upon the original alpha level when comparing multiple means, which increases the probability of error (Bruning and Kintz 1971). Had no significant differences been detected, a less conservative method of comparing means, such as, Tukey's Honestly Significant Difference would have been used. Protecting the critical alpha level guarded against a type II error occurring during data analysis.

This study both supported and contradicted the 1988 study by O'Conner, Bahrke, and Tetu. The results of this study supported their study in that age category 1 (17-21 years) soldiers generally scored the lowest of the age categories in every event except the 2-mile run. As a group, they also exhibited the most failures of any age category. Also, age categories 4 (32-36 years) and 5 (37-41 years) generally scored the highest of all age categories in every event, except the push-up event.

O'Conner et al concluded that soldiers in the youngest age category scored the lowest due to a lack of proper physical training supervision and motivation, not indicating that the standards may have been too high.

The results of this study contradicted the O'Conner et al study in that age category 3 (27-31 years) soldiers scored the lowest of any age category on the 2-mile run. Also, failure rates on every event, except the push-up, increased at age categories 6 (42-46 years) and 7 (47-51 years) rather than continuing the downward trend.

This study also supported and contradicted the 1994 study by Kahue, Reed, Kiyokawa, and Finstuen. The results of this study supported that study in that age category 1 (17-21 years) soldiers generally scored the lowest of the age categories on the APFT total score. Also, no statistical significance was found for the main effect of gender means on the grand mean APFT total score as well as no statistically significant interaction effects for gender and age category were observed.

This study contradicted the Kahue et al study in that no statistically significant results were obtained in the test of main effect differences for age category means on the grand mean APFT total score. However, statistical significance was obtained for the main effect of age category means and the grand mean APFT 2-mile run score. Also, in the Kahue et al study, age category 1 (17-21 years) soldiers exhibited the smallest standard deviation scores. Whereas, in this study, the dispersion of their scores

were the largest among the age categories for the APFT total score.

The outcome of this study may have had several contributing factors. First, soldiers in age categories 4 (32-36 years) and 5 (37-41 years) are mid- and late-career soldiers. Higher scores from these soldiers may reflect their realization of the importance of performing well on the APFT for retention and promotion purposes. Second, soldiers in age category 3 (27-31 years) scored relatively well in the push-up and sit-up events, but did very poorly as a group on the 2-mile run. Many interns and residents fall into this age category along with nursing students. The long hours of medical training may detract from their ability to aerobically train for the run. It was also interesting to note that the number of maximum scores in the push-up and sit-up events were unremarkable across the age categories. However, in the 2-mile run event, age categories 1 (17-21 years) and 3 (27-31 years) each had only 1 test subject maximize the event. All other age categories ranged from 5 to 8 test subjects maximizing the 2-mile run event.

There were several limitations to this study. Although strict standards for grading the APFT are outlined in FM 21-20, it is realized that differences may exist in judgment for graders as to what constitutes a correct repetition of an exercise. Also, this study did not contain data for all APFT age categories. Insufficient numbers of subjects in age category 8 (52+ years), precluded testing of this age category for which

test standards are currently set. Finally, verification of scores were not possible because of the retrospective study design.

5. CONCLUSIONS AND RECOMMENDATIONS

Conclusions of this study are warranted utilizing a retrospective data analysis, experimental design. Randomly selected samples were used in a two-factorial analysis of variance, and mean comparisons were tested for significance using Scheffe's method of comparing multiple means.

Future research should focus on a larger sample size that encompasses all Military Occupational Skills and Areas of Concentration for the Army. This will give the U.S. Army Training and Doctrine Command, who approves physical fitness standards, a better indication of physical performance standards for the entire force.

There are several implications of this study. First, standards that are not fair across all age and gender categories affect reenlistment options of enlisted soldiers, since they must pass the APFT in order to reenlist. Second, a soldier's chance of promotion is affected since the APFT score can generate up to 50 promotion points for junior enlisted soldiers, and noncommissioned and commissioned officers must pass the APFT to receive a promotion. Finally, nursing students and post-graduate medical personnel must pass the APFT to complete their medical training programs.

Retention and promotion of medically trained soldiers are

important since the skills attained by these soldiers are in high demand in the civilian sector. The training they receive is expensive and lengthy. The time and money required to retrain new personnel may be utilized for other requirements. Unlike other soldiers, the soldiers in the medical professions are not interchangeable.

Findings from this study may warrant further evaluation since APFT failures may result in removal of soldiers from schools and repeated failures result in separation from the Army.

APPENDIX A - DEFINITION OF TERMS

1. Alpha Level (α) - The level of significance for testing hypotheses. It is the probability value researchers use to decide whether or not to reject the null hypothesis. There is no single standard or universal level. However, the higher the significance level used for testing a hypothesis, the higher the probability of rejecting a null hypothesis when it is true. Scientific researchers normally set this level at a minimum of .05. This means that the probability that the results obtained are due to chance alone are less than 5 percent. This 5 percent, is the area that falls outside of the 95 percent confidence limit of a normally and independently distributed sample.
2. Analysis of Variance (ANOVA) - A statistical procedure that determines whether or not there are any differences among two or more groups of subjects on one or more factors.
3. Interaction - a relationship between two independent variables such that they have a different effect on the dependent variable. Interaction occurs when the effect of one factor depends on the level of the other factor. The presence of interaction implies that Factor A and Factor B both have effects on the dependent variable.
4. Main Effect - An estimate of the effect of an independent variable measured separately of other independent variables which may form part of an experiment. The term does not imply causality. The main effect of factor A is the average of the estimates of the effect of Factor A when Factor B is held constant at its higher and lower level.
5. Type II Error - Accepting a null hypothesis when it is false. This can be very damaging in healthcare research, because a truly statistically significant event is missed.

(Snedecor and Cochran 1967)

APPENDIX B - TABLES OF APFT FAILURE RATES IN INITIAL AND FINAL STUDY SAMPLES

TABLE 9. APFT Failure Rates in Initial Study Sample

Age Category	# Male	# Fail	%	# Female	# Fail	%	Total #	# Fail	%
1	40	7	17.50	13	4	30.76	53	11	20.75
2	114	8	7.01	53	6	11.32	167	13	7.78
3	114	6	5.26	64	7	10.94	178	13	7.30
4	126	4	3.17	43	1	2.33	169	5	2.96
5	89	2	2.25	39	2	5.13	128	4	3.13
6	32	0	0.00	25	2	8.00	57	2	3.51
7	25	0	0.00	11	1	9.09	36	1	2.78
Overall	540	27	5.00	248	23	9.27	788	49	6.22

TABLE 10. APFT Failure Rates for Final Study Sample

Age Category	# Male	# Fail	%	# Female	# Fail	%	Total #	# Fail	%
1	11	2	18.18	11	3	27.27	22	5	22.73
2	11	1	9.09	11	1	9.09	22	2	9.09
3	11	2	18.18	11	1	9.09	22	3	13.64
4	11	0	0.00	11	0	0.00	22	0	0.00
5	11	0	0.00	11	0	0.00	22	0	0.00
6	11	0	0.00	11	1	9.09	22	1	4.55
7	11	0	0.00	11	2	18.18	22	2	9.09
Overall	77	5	6.49	77	8	10.39	154	13	8.44

APPENDIX C - SUMMARY OF SUPPORT/NONSUPPORT OF HYPOTHESES

H1) Accept the null.

Ho: There are no significant main effect differences for age category and grand mean APFT total score.

H2) Accept the null.

Ho: There are no significant main effect differences for gender and grand mean APFT total score.

H3) Accept the null.

Ho: There are no interaction effects between age category and gender in the grand mean APFT total score.

H4) Accept the null.

Ho: There are no main effect differences for age category and grand mean APFT push-up score.

H5) Accept the null.

Ho: There are no main effect differences for gender and grand mean APFT push-up score.

H6) Accept the null.

Ho: There are no interaction effects between age category and gender in the grand mean APFT push-up score.

H7) Accept the null.

Ho: There are no main effect differences for age category and grand mean APFT sit-up score.

H8) Accept the null.

Ho: There are no main effect differences for gender and grand mean APFT sit-up score.

H9) Accept the null.

Ho: There are no interaction effects between age category and gender in the grand mean APFT sit-up score.

H10) Reject the null, and accept the alternate hypothesis.

Ha: There are main effect differences for age category and grand mean APFT 2-mile run score.

H11) Accept the null.

Ho: There are no main effect differences for gender and grand mean APFT 2-mile run score.

H12) Accept the null.

Ho: There are no interaction effects between age category and gender in the grand mean APFT 2-mile run score.

APPENDIX D - SCHEFFE'S FORMULA FOR COMPARING MULTIPLE MEANS AND CALCULATIONS FOR SIGNIFICANCE

Scheffe's Formula:

$$\text{Critical Value} = \sqrt{(a-1)F_{0.05, a-1, a(n-1)}} \times \sqrt{2MSe/n}$$

APFT 2-Mile Run Critical Value:

$$a = 7, n = 22, MSe = 432.11$$

$$\text{Critical Value} = \sqrt{(7-1)2.10} \times \sqrt{2(432.11)/22}$$

$$\text{Critical Value} = 22.25$$

Significance at the .05 alpha level is achieved if the absolute value of the difference of mean comparisons exceeds the critical value.

$$\text{Age Category 1} - \text{Age Category 2} = 69.32 - 76.09 = 6.77$$

$$\text{Age Category 1} - \text{Age Category 3} = 69.32 - 63.86 = 5.46$$

$$\text{Age Category 1} - \text{Age Category 4} = 69.32 - 87.68 = 18.36$$

$$\text{Age Category 1} - \text{Age Category 5} = 69.32 - 87.50 = 18.18$$

$$\text{Age Category 1} - \text{Age Category 6} = 69.32 - 81.68 = 12.36$$

$$\text{Age Category 1} - \text{Age Category 7} = 69.32 - 76.27 = 6.95$$

$$\text{Age Category 2} - \text{Age Category 3} = 76.09 - 63.86 = 12.23$$

$$\text{Age Category 2} - \text{Age Category 4} = 76.09 - 87.68 = 11.59$$

$$\text{Age Category 2} - \text{Age Category 5} = 76.09 - 87.50 = 11.41$$

$$\text{Age Category 2} - \text{Age Category 6} = 76.09 - 81.68 = 5.59$$

$$\text{Age Category 2} - \text{Age Category 7} = 76.09 - 76.27 = 0.18$$

$$\text{Age Category 3} - \text{Age Category 4} = 63.86 - 87.68 = \mathbf{23.82}$$

$$\text{Age Category 3} - \text{Age Category 5} = 63.86 - 87.50 = \mathbf{23.64}$$

Age Category 3 - Age Category 6 = 63.86 - 81.68 = 17.82

Age Category 3 - Age Category 7 = 63.86 - 76.27 = 12.41

Age Category 4 - Age Category 5 = 87.68 - 87.50 = 0.18

Age Category 4 - Age Category 6 = 87.68 - 81.68 = 6.00

Age Category 4 - Age Category 7 = 87.68 - 76.27 = 11.41

Age Category 5 - Age Category 6 = 87.50 - 81.68 = 5.82

Age Category 5 - Age Category 7 = 87.50 - 76.27 = 11.23

Age Category 6 - Age Category 7 = 81.68 - 76.27 = 5.41

REFERENCE LIST

- Bielenda, Christine C., Knapik, Joseph, and Wright, David. 1993. Physical Fitness and Cardiovascular Disease of Female Senior U.S. Military Officers and Federal Employees. Military Medicine, 158(3):177-81.
- Bruning, James L. and Kintz, B. L. Computational Handbook of Statistics, 2d ed., Glenville, Illinois: Scott, Foresman and Company, 1971.
- Cline, William. Good Training Helps 'Old Soldiers' Pass PT Test. *The Mercury*, February 1994.
- Dowdy, Shirley and Wearden, Stanley, Statistics for Research, 2d ed., New York: John Wiley and Sons, 1991.
- Headquarters, Department of the Army, Washington, D.C., Physical Fitness Training, Field Manual 21-20, September 30, 1992.
- Headquarters, Department of the Army, Washington, D.C., The Army Physical Fitness Program, AR 350-15, December 1989.
- Headquarters, Department of the Army, Washington, D.C., Officer Evaluation Reporting System (with changes), AR 623-105, March 31, 1992.
- Headquarters, Department of the Army, Washington, D.C., Enlisted Evaluation Reporting System (with changes), AR 623-205, March 31, 1992.
- Kahue, Philip, Reed, Bruce R., Kiyokawa, Guy, and Finstuen, Kenn. "Mean Comparisons of Age and Gender Differences on the Total Score of the Army Physical Fitness Test (APFT)" Ft. Sam Houston, Texas: AMEDD Center and School.

- Knapik, Joseph. 1989. The Army Physical Fitness Test (APFT): A Review of the Literature. Military Medicine, 154(6):326-9.
- Knapik, Joseph, Ang, Philip, Reynolds, Katy, and Jones, Bruce. 1993. Physical Fitness, Age, and Injury Incidence in Infantry Soldiers. Journal of Occupational Medicine, 35(6):598-603.
- O'Connor, John S., Bahrke, Michael S., and Tetu, Robert G. 1990. 1988 Active Army Physical Fitness Survey. Military Medicine, 155(12):579-85.
- Pollock, ML, Wilmore, JH, Fox, SM. Exercise in Health and Disease. Philadelphia, W.B. Saunders Company, 1984.
- Sharkey, BJ. Physiology of Fitness. Champaign, Illinois. Human Kinetics Publishers Inc., 1984.
- Snedecor, George W. and Cochran, William G. Statistical Methods. 6th ed., Ames, Iowa: Iowa State University Press, 1967.
- Song, Thomas M., and Moore, John. 1989. Physical Fitness of Militia Forces. Military Medicine, 154(9):477-9.